DOCUMENT RESUME

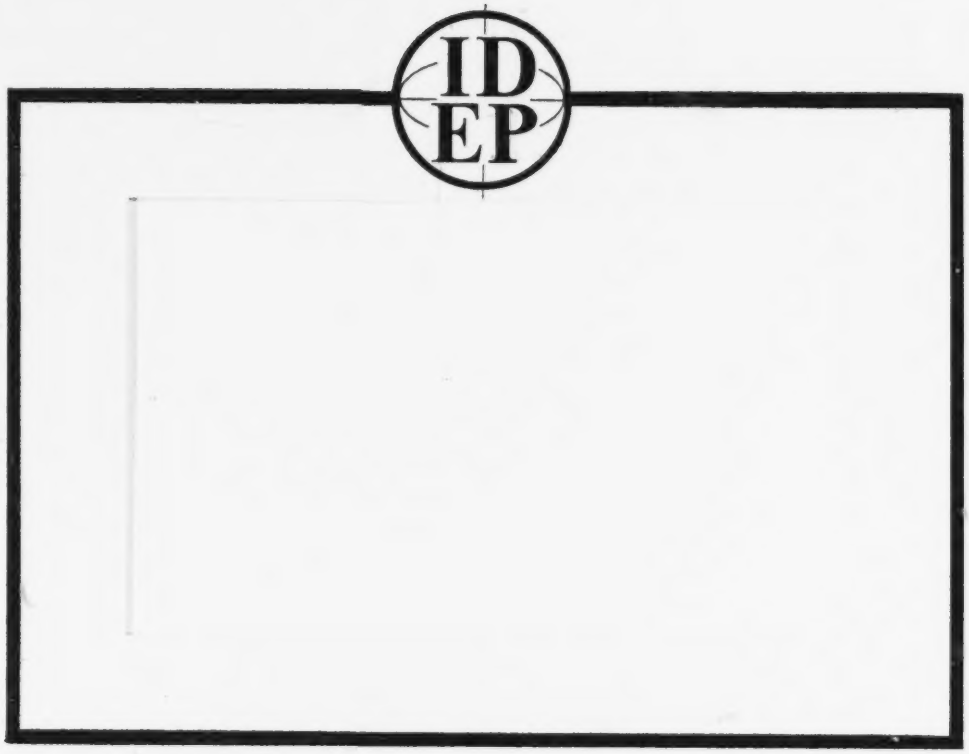
ED 071 946 SO 005 140

AUTHOR	Tate, Sean Alan
TITLE	Evaluation in the Social Studies: A Critical Review of the Literature.
INSTITUTION	Pittsburgh Univ., Pa. International and Development Education Program.
PUB DATE	Oct 71
NOTE	65p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Comparative Education; *Curriculum Evaluation; Evaluation Methods; *Evaluation Needs; Literature Reviews; *Social Studies; State of the Art Reviews

ABSTRACT

This paper looks at curriculum evaluation as it operates today and proceeds to examine evaluation in the social studies as viewed by some of the well-known authors in the field. This is followed by a review of evaluation used by some of the various social studies curriculum projects and presents a number of findings. The paper, according to the author, could be seen as a kind of informal "meta-evaluation" i.e., the evaluation of evaluation. The paper concludes that a formal meta-evaluation of social studies curriculum projects is needed if the field is to catch up to other disciplines and become more cohesive and worthwhile. Two appendices are given: the first contains a listing of curriculum development projects in social studies, while the second contains some definitions of the term "evaluation". A seven page bibliography is also found at the end of the paper. (FDI)

ED 071946



SP005140

INTERNATIONAL AND DEVELOPMENT EDUCATION PROGRAM

SCHOOL OF EDUCATION

UNIVERSITY OF PITTSBURGH

ED 071946

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

EVALUATION IN THE SOCIAL STUDIES: A CRITICAL REVIEW OF THE LITERATURE

by

Sean Alan Tate

International and Development Education Program
School of Education
University of Pittsburgh
Pittsburgh, Pennsylvania 15213

October 1971

If we do not strive to enrich our traditional fragmentary evaluations of selected aspects of education by continuing reference to some notion of wholeness of impact, however vague and ambiguous, our evaluations are likely to remain at the level of part-time bookkeeping.

Melvin M. Tumin

TABLE OF CONTENTS

	Page
INTRODUCTION	1
CURRICULUM EVALUATION: A GENERAL SURVEY	3
Defining Curriculum Evaluation	3
Methodological Approaches	5
Emphases	6
Some Distinctions	7
Problems in Curriculum Evaluation	9
Conclusions	14
SOCIAL STUDIES EVALUATION: A REVIEW OF THE LITERATURE	15
The Problems	15
Social Studies Evaluation Models	18
Evaluation of Social Studies Materials	21
Conclusions	23
EVALUATION AND THE SOCIAL STUDIES PROJECTS	24
Evaluation Problems in Curriculum Projects	24
Evaluation in Social Studies Projects: A Review	25
Conclusions	32
EVALUATION IN THE SOCIAL STUDIES: THE INTERNATIONAL PERSPECTIVE	33
International Organizations and Social Studies Evaluation	33
Social Studies Evaluation in Other Nations	34
CONCLUSION	37
APPENDIX I: Curriculum Development Projects: Social Studies Education	39
APPENDIX II: Evaluation: Some Definitions	45
BIBLIOGRAPHY	55

INTRODUCTION

"Social Studies," as practiced in American schools has undergone a slow but steady transformation in the last ten years. Traditionally Social Studies has been equated with dull and dry excursions into only two major areas of concern: history and geography. The student was required to memorize vast mountains of facts and dates, usually retrieved from textbooks as dull as the courses themselves. Although curriculum development in the Social Studies continues to lag behing development in Science, Reading, and Mathematics, some important strides are nevertheless being made. These are, for example, now over 40 different curriculum projects in operation in the United States which are concerned with Social Studies. Even more importantly, the term "Social Studies" has been considerably broadened out of its old history-geography base to include anthropology, economics, political science, and sociology. In addition to the textbooks which have been written to deal with each of the above areas, Social Studies curriculum developers have been gradually retreating from the use of textbooks as the sole tool for Social Studies instruction. Now slides, movies, records, documents, groups of readings, and field studies comprise an ever more important part of the curriculum. Despite all the innovations in the "New Social Studies" (a term initiated by Edwin Fenton of Carnegie-Mellon University), the use of new methods, such as the inquiry method, still lags far behind, in large part because teachers are not effectively trained to use the materials and methods involved.

With the advent of increasing interest in Social Studies curriculum and the funding of a large number of Social Studies curriculum projects, the problem of evaluation has arisen. In many cases the term seems to have left some curriculum developers floundering. The frustration is evident in the words of Terry Denny, who, as editor of an issue of the Educational Product Report (February 1969) on educational evaluation, began his introduction by stating that "evaluation is a worrisome word in educational parlance that resists definition about as stoutly as any concept in vogue."[1] An inkling of the number and variety of definitions of "evaluation" can be seen from a brief perusal of Appendix I of this paper which brings together a few of these definitions. Current uses of educational evaluation include improving on-going programs, rating one project against another, assessing the merit of a terminated activity, providing counsel for work about to begin, seeking valid principles for generalization to other similar efforts and recording fully the story of a particular educational effort. From these uses

---

[1] Terry Denny. Educational Product Report, Vol. 2, No. 2 (February 1969).

and others, it would be difficult to say that there is "one" definition of evaluation which is "best" to use, and it is not fruitful to try and do so. Although it is informative to examine the other definitions, it appears that definitions of evaluation are "good" in different contexts.

This paper will look briefly at curriculum evaluation as it operates today, will proceed to examine evaluation in the Social Studies as viewed by some of the well-known authors in the field, then review evaluation as used by some of the various Social Studies curriculum projects, and present a number of findings. The paper could be seen as a kind of informal "meta-evaluation," a term coined by Michael Scriven to mean "second-order evaluation, i.e., the evaluation of evaluation."

> Theoretically, meta-evaluation involves the methodological assessment of the role of evaluation; practically, it is concerned with the evaluation of specific evaluative performances.[2]

What is really needed at this point, however, is formal meta-evaluation of curriculum projects' evaluation programs, not just in Social Studies, but in all the other areas of curriculum development where evaluation is being used.

Many of the materials used in this paper came from the Social Studies Curriculum Center of Carnegie-Mellon University, and gratitude must be expressed to the two secretaries of the Center, Mrs. Ethel Strasser and Catherine Dudas, for their help and patience. Carol Jones of the International and Development Education Program, University of Pittsburgh, typed the manuscript.

---

[2] Michael Scriven. "An Introduction to Meta-Evaluation." Educational Product Report, Vol. 2, p. 36 (February 1969).

CURRICULUM EVALUATION: A GENERAL SURVEY

Defining Curriculum Evaluation

"Curriculum evaluation" is a term which, by its range of definitions, is about as vague as a bank of fog. The component parts of the term are equally nebulous. The word "curriculum" in its traditional sense conjures up such definitions as "course of study" but in fact curriculum now seems to have a wide range of meanings, particularly with the recent push by such people as John Holt, Charles Silberman, Neil Postman, Charles Weingartner, and Ivan Illich against the traditional school and its methods. Definitions of curriculum have now become very broad. Wayne Welch notes that:

> Curriculum is frequently defined by educational theorists as the link between society and the schools, the major source of stimuli found in instructional settings, or as pupil behavior pertinent to the goals of the school. According to these definitions, curriculum could include everything that happens to a child in school from learning how to accept last Friday's football game loss to the complete sequence of experiences that produced his 800 SAT score.[1]

Curriculum is therefore a rapidly changing term and field of study and is now going beyond what happens in the school to include experiences exterior to the school environment as well. And because the term is broad and in a state of flux, it is little wonder that people have difficulty in agreeing in what "curriculum" is. Another example of the breadth in the definition of curriculum may be seen in the definition presented by Thomas, Sands, and Brubaker that:

> Curriculum means all the intended learning goals, experiences, teaching materials and evaluation techniques which educators plan and/or use [2]

Definitions of evaluation also range from the narrowest of focuses to extremely broad visions. (A short compilation of some of these definitions appears in Appendix I of this paper.) Melvin Tumin states that:

> Evaluation means many things to many different people, not only because they are defensive about the possible

---

[1] Wayne Welch. "Curriculum Evaluation." Review of Educational Research, Vol. 39, No. 4 (1969), pp. 429-443.

[2] R. Murray Thomas, Lester Sands, and Dale Brubaker. Strategies of Curriculum Change. Scranton, Pa.: International Textbook Company, 1968, p. 6.

> results of a systematic scrutiny of their effectiveness, but because too they have different notions as to what ought to be transpiring in any educational transaction and what are the legitimate sources of pride and shame. So too they have different criteria... criteria as to what constitutes evidence... and they often cannot see how a non-participant can possibly evaluate as well as a participant in an enterprise.[3]

Curriculum evaluation, then, is and has been just as illustrative as its component terms. Ian Westbury, in a recent and thorough review of curriculum evaluation, finds that:

> ...the assertion that "we must evaluate our curricula" in terms of cost, effectiveness, content and the like has a ring of sense and efficiency and a commonplace obviousness that makes it impossible to believe the opposite. To this extent everyone supports the evaluation of curricula. Curriculum evaluation is, however, another thing: it is a body of techniques, methodologies, and principles created deliberately (and recently) to give some systematic form to the ways in which the assertion "we must evaluate"... can be made to work.[4]

Ronald C. Doll, in the second edition of his well-known work on Course Improvement, defines evaluation as:

> ...a broad and continuous effort to inquire into the effects of utilizing educational content and process according to clearly defined goals.[5]

It can already be seen that the definitions of curriculum evaluation vary from author to author. Such well-known names in the field of educational evaluation as Guba, Stake, Stufflebeam, Cronbach, and Forehand all make their contributions. The late Hilda Taba perhaps best outlined the program of curriculum evaluation when she noted that the term evaluation covers an extraordinary variety of meanings and processes. For example, one can evaluate anything about the schools' curriculum: its objectives, its scope, the quality of personnel in charge of it, the capacities of students, the relative importance of various subjects, the degree to which objectives are implemented, and the effectiveness of the equipment and materials.

---

[3] Melvin Tumin. "Evaluation of the Effectiveness of Education." Interchange, Vol. 1, No. 3 (1970), p. 106.

[4] Ian Westbury. "Curriculum Evaluation." Review of Educational Research, 4: 240 (April 1970).

[5] Ronald C. Doll. Course Improvement. (Boston: Allyn and Bacon, 1970), p. 379.

Secondly, Taba noted that evaluation may be used to refer to different processes from a rendering of a value judgement based on opinion to "the careful gathering of evidence on the attainment of objectives, a forming of judgements on the basis of that evidence, and the weighing of that evidence in the light of those objectives."[6] Thirdly, evaluation can be carried on at a variety of levels and by different categories of people.

The narrowest concept of evaluation, Taba stated, defined evaluation as marking or grading. The wider definition involves the process which includes the formulation of objectives, decisions about the means of securing evidence, processes of interpretation, and decisions about the needed changes and improvements in curriculum.[7] Another point of confusion in curriculum evaluation is the identification of evaluation with measurement. Many authors write a book entitled "Evaluation..." and after several introductory chapters they wander into discussion of the varieties of measurement (intelligence tests, etc.). Measurement, of course, is an important part of evaluation, but it definitely should not be equated with the larger concern of "valuing."

Methodological Approaches

A variety of methodological approaches are used in curriculum evaluation. More often than not evaluations are of an informal nature, particularly at the local level. Informal evaluation of curricula has a number of drawbacks, not notably its lack of objective and comprehensive judgement criteria. The generalizability of informal and curricula evaluation is accordingly an important and continuing problem.

What has been described as a formal evaluation of curricula is exemplified by various works of Cronbach, Stevens and Morrisset, and others. This type of evaluation often does set up criteria and objectives and can include the decision-making process within its domain. The use of formal evaluation is most often found in large curriculum projects and offers greater generalizability over a wide spectrum.

Another approach to curriculum evaluation is that of program accounting, or cost/benefit analysis. A prime example is PPBS (Planning, Programming, Budgeting System) which, while not strictly an evaluative device, is thought of as a "system for relating inputs to outputs in such a way as to most effectively allocate resources in relation to objectives." A PPB System is

---

[6] Hilda Taba. Curriculum Development. (New York: Harcourt, Brace, and World, 1962), p. 310.

[7] Ibid.

a management and planning tool (rather than an accounting tool) for "more efficient use of limited resources in the achievement of specified institutional goals." The basic concepts of PPBS have been summarized by Freeman:

1. Explicit identification of objectives based on comparisons and analyses of alternative objectives.

2. Systematic examination of alternative means of accomplishing given objectives.

3. Explicit identification of total costs of each alternative, in terms of real costs (money, space, time, manpower) and long-term costs of current decisions.

4. Multi-year planning and programming system.

5. Organization of budgets on the bases of objective-oriented programs.

6. Explicit comparisons of costs and benefits of each program.[8]

PPBS is therefore not purely an evaluation system per se; and while it has achieved wide support from a number of sectors (particularly State and Federal governments), it has also been the target for sniping activity from those who feel that cost accountability and other such quantifiable aspects ignore too many important but non-quantifiable factors.

Emphases

Curriculum evaluation has seen a number of different emphases, depending on who has developed the models. Stufflebeam, for example, while not specifically working with just evaluation of curricula, emphasizes the neglected area of decision-making in many of his studies. His model for educational evaluation includes four different components: i.e., context evaluation, input evaluation, process evaluation, and product evaluation (CIPP).[9]

Another emphasis in curriculum evaluation manifests itself in the evaluation of different aspects of the curriculum entity. A prime example of this type is evaluation of curriculum materials, such as textbooks, workbooks, and teachers' manuals. Another example in this realm is the

---

[8] Jack Freeman. PPBS. (Paper presented at the University of Pittsburgh, 1970), p. 1.

[9] Daniel Stufflebeam and Egon Guba. Evaluation: The Process of Stimulating, Aiding, and Abetting Insightful Action. (Columbus: Evaluation Center, Ohio State University, 1968), p. 50.

evaluation of various pieces of educational equipment such as tape recorders, motion-picture projectors, and teaching machines. Perhaps the best-known example is a rather expensive journal called Educational Product Report, a publication of the Educational Products Information Exchange, which provides "descriptive and evaluative information about all types of learning materials, equipment, and systems." The information, however, tends to be more descriptive than evaluative. The journal provides many convenient tables which allow its readers to compare specifications and costs of various pieces of equipment, texts, and other materials, but the tables seldom provide any comparative judgments about these items. Occasionally separate articles do examine specific curriculum materials in a more thorough manner. The descriptive Stevens-Morrisset System has often been employed in these analyses. The Report has, from time to time, devoted itself to the problems of evaluation. Maurice J. Eash, for example, authored an article entitled, "Assessing Curriculum Materials: A Preliminary Instrument." Eash's instrument has the advantage of being concise and easy to use but like the Stevens-Morrisset system its product would appear to be mostly description.[10]

Some Distinctions

Several widely recognized distinctions in curriculum evaluation might be examined at this point. One is Michael Scriven's "formative-summative" division. Formative evaluation indicates evaluation of a program still in progress, while summative evaluation refers to an evaluation at the end of a program. Another distinction is that between description and evaluation. Arlene Payne makes this distinction between analyzing curricula to describe them, without reference to standards of what they should be, and to evaluate them by comparing them with a model or criteria. This distinction is most helpful because it appears that a good deal of what is purported to be "evaluation" is in reality only "description."[11]

A third distinction is that between curriculum evaluation and curriculum analysis made by Alan Tom. He refers to the examination of the worthiness of curricular contents as curriculum analysis and uses the term curriculum evaluation when the purpose is to discover whether a curriculum achieves what it purports to achieve. However a number of writers continue to use the term analysis to mean evaluation. Tom feels that the distinction is an important one and uses a simple analogy to clarify the difference between the two processes:

---

[10] Maurice J. Eash. "Assessing Curriculum Materials: A Preliminary Instrument." Educational Product Report, 2:18-24 (February 1969).

[11] Arlene Payne. The Study of Curriculum Plans. (Washington, D. C.: National Education Association, 1969).

> Suppose we want to decide whether to use a particular insecticide. One question we would ask is whether an insecticide does what it is supposed to do, i.e., destroy insects. Yet the questions of whether we should use the insecticide is a broader problem. The latter question raises such issues as: special equipment needed to spread the insecticide, the effect of the insecticide on wildlife, the cost of using the insecticide and the effect, if any, on humans.[12]

Tom notes that the distinction has also been recognized by several other writers. Michael Scriven, for example, has argued that "if the goals aren't worth achieving, then it is uninteresting how well they are achieved..."[13] Thus evaluation proper must include, as an equal partner with the measuring of performance agains goals, procedures for the evaluation of goals.

Still another distinction by Melvin Tumin is that of the "central polarity in the field of evaluation today, i.e., the dialectic between wholeness and fragmentation." Tumin sees a struggle in the evaluation field

> between trivial precision and apparently rich ambiguity, and it is imperative that we strike a better balance between these than has been true in the past.[14]

Tumin wants evaluator "to strive to enrich our traditional fragmentary evaluations of selected aspects of education by continuing reference to wholeness of impact."[15] However, he also sees the practical limitations of the holists' views (the holists being those who look down their noses at such things as measures, controls, samples, indicators, and the rest of the paraphernalia of scientific evaluation). He finds that most of the important questions in evaluation cannot be answered unless the evaluator is willing to "accept very fragmented and partial indicators as somehow standing for the whole." The problem, he states, is a result of unavoidable disjunction in the human mind between big words and thoughts."[16]

---

[12] Alan Tom. An Approach to Selecting Among Social Science Curricula. (Paper presented at Metropolitan St. Louis Social Studies Center, 1970), p. 61.

[13] Michael Scriven. Perspectives of Curriculum Evaluation. (Chicago: Rand McNally [AERA Monograph 1], 1967), p. 52.

[14] Melvin Tumin. "Evaluation of the Effectiveness of Education." Interchange, 1:98 (1970).

[15] Ibid.

[16] Ibid.

Problems in Curriculum Evaluation

Having looked at these four distinctions, or core problems, we may now look at some of the other problems which curriculum evaluation faces. A major problem, one of definition, comes under the heading of evaluation design. Hawkridge, for example, defines "design" as "the overall plan of the evaluation and the strategies within each phase of the plan." He identifies seven phases in educational evaluation:

1. Setting up objectives for evaluation.
2. Selecting objectives to be measured.
3. Choosing instruments and procedures.
4. Selecting samples.
5. Establishing measurement and observation schedules.
6. Choosing analysis techniques.
7. Drawing conclusions and recommendations.[17]

These seven components, while not necessarily all-inclusive, bring to mind a number of additional problems. There is, for example, the problem of objectives. Gagné and Mager feel that objectives must be precise, detailed descriptions of student behavior exhibited or the attainment of an objective. The whole argument on objectives centers on their specificity. Many educators feel that behavioral objectives constrict education to "trivial kinds of behavior that can be described precisely." These critics feel that the teacher's spontaneity in the classroom may be threatened by such objectives. Bloom makes a more moderate view on this problem. He suggests that:

> ...it is virtually impossible to engage in an educational enterprise of any duration without some specification to guide one. Insofar as possible, the purpose of education and the specification of educational changes should be made

---

[17] David Hawkridge. "Designs in Evaluative Studies." in Evaluative Research. (Pittsburgh: American Institutes for Research, 1970), p. 27.

> explicit if they are to remain open to inquiry, if teaching and learning are to be modified or change is needed, and if each new group of students is to be subjected to a particular set of educational processes.[18]

Hawkridge notes that the setting up of objectives for an evaluation are not necessarily the same as setting up objectives for a program, especially since the objectives for a program are often set up by somebody other than the evaluator.

Guba and Stufflebeam have carefully examined one type of design... experimental design...and have described the problems that may arise when it is applied to evaluations. Experimental designs would seem to be reasonable methods for evaluation problems "since traditionally both experimental research and evaluation have been used to test hypotheses about the effects of treatments." Guba and Stufflebeam, however, see four flaws with this reasoning:

1. First, the application of experimental design to evaluation problems conflicts with the principle that evaluation should facilitate the continual improvement of a program.
   Experimental design prevents rather than promotes changes in the treatment because treatments cannot be altered in process if the data about differences between treatments are to be unequivocal.

2. A second flaw in experimental design type of evaluation is that it is useful for making decisions after a project has run full cycle but almost useless as a device for making decisions during the planning and implementation of the project.

3. A third problem with the experimental design type of evaluation is that it is suited to the antiseptic conditions of the classroom.

4. A fourth flaw inherent in the application of conventional experimental design is the possibility that while internal validity may be gained through the control of extraneous variables, such an achievement is accomplished at the expense of external validity.[19]

---

[18] Benjamin Bloom. "Some Theoretical Issues Relating to Educational Evaluation." in Educational Evaluation: New Roles, New Means. (Chicago: Sixty-Eighth Yearbook, National Society for the Study of Education, 1969).

[19] Stufflebeam and Guba, op. cit., p. 14.

The problem of defining criteria for judging the worth of evaluations is also of major importance for if inappropriate or insufficient criteria are used, the result may be faulty designs and useless reports.

In sum, Guba and Stufflebeam have attempted to summarize the problems of educational evaluation, including curriculum evaluation, as follows:

1. Inadequacies of present definitions.

2. A lack of understanding of the different educational settings within which evaluation must be conducted.

3. A lack of understanding of generalizable information requirements which educational evaluation studies must meet.

4. The lack of valid structure for the generalizable parts of the evaluation design.

5. The lack of an appropriate set of criteria for judging the worth of evaluation strategies, designs, instruments, reports, etc.

6. The lack of concepts needed to organize and operate evaluation systems.[20]

Several other writers in the field of educational evaluation and particularly curriculum evaluation have also suggested the same problems plus a number of others. For example, Herbert H. Walberg, looking specifically at "Curriculum Evaluation: Problems and Guidelines," has noted the previously mentioned controversy on objectives and the problems posed by the educational environment. Walberg carefully examines a problem stated in Guba and Stufflebeam's list: the problem of generalizability:

> That evaluation should be generalizable to specified populations of students seems an obvious objective; yet most evaluations must be faulted on statistical grounds.[21]

---

[20] Ibid.

[21] Herbert J. Walberg. "Curriculum Evaluation: Problems and Guidelines." Teachers College Record, 71:564 (May 1970).

Walberg faults curriculum evaluation statistically by its lack of random sampling, thereby limiting inferences to the population from which the sample has been drawn.

A second part of the generalizability problem involves the changes in a course and the students across time. Walberg asks: "To what extent does a course remain unchanged while undergoing evaluation?"[22] Formative evaluation is, of course, helpful here, but even at the stage of summative evaluation, the course may still be evolving and yearly cycles of summative evaluation may be needed for a few years after the course is completed.

A third problem involving generalizability across time is "the changing state of society and the possible irrelevance of courses developed before relevant changes."[23]

Two other writers have faced the generalizability problem in the Educational Products Report: Robert Stake and James L. Wardrop. Stake sees two approaches to educational evaluation. He believes there is a choice between being (1) scientific, generalizing and evaluating to find out why, or (2) to be descriptive, to be delimited, and to evaluate to find out what.[24] At times Stake thinks that the evaluator should opt for evaluation outside the "scientific process, depending on the evaluative job." He refers to evaluation that permits generalization in many directions as "higher evaluation," while "lower evaluation" yields conclusions limited to a specific setting."

Wardrop echoes Stake's distinction between evaluation for the "whats" and "whys." He also notes that

> ...whether or not an evaluation study is designed for generalizability (to other classrooms, other children, other communities or other times), the consumer (sponsor, participant, or interested laymen) will make generalizations from its results, even though such generalizations may be unwarranted or even illegitimate. The evaluator cannot control the ways in which others use the results of his evaluation study.[25]

---

[22] Ibid.

[23] Ibid.

[24] Robert Stake. "Generalizability of Program Evaluation." Educational Product Report 2:41 (February 1969).

[25] James Wardrop. "Generalizability: The Dangers of Limits." Educational Product Report 2:42 (February 1969).

Wardrop sees the evaluators' most effective safeguard on this problem as the use of the most scientific and most generalizable design possible.

Egon Guba has presented one of the best critiques on the flaws of educational evaluation in an article entitled, "The Failure of Educational Evaluation." While not strictly about curriculum evaluation, Guba finds that the basic lacks of educational evaluation today are numerous and they can be quite easily applied to curriculum evaluation:

1. Lack of adequate definition of evaluation,
2. Lack of adequate evaluation theory,
3. Lack of knowledge about decisions processes,
4. Lack of criteria,
5. Lack of approaches differentiated by levels,
6. Lack of mechanisms for organizing processing and reporting evaluative information.[26]

Many of the problems voiced by Stufflebeam, Stake, Wardrop, Guba, and Walberg come down to the operationalization of an evaluation program or evaluation model. This is, of course, the problem with most models: they look good on paper but when it comes to putting them in operation, they sometimes fall flat. Many models also lack comprehensiveness, while others totally ignore the decision-making area. Barbara Cass,[27] after examining seven different models, finds that the Stufflebeam model (CIPP) is strongest in this regard.[28] A good example of a model which is apparently difficult to operationalize is Malcom Provus' Discrepancy Model which has been used in the Pittsburgh Public Schools and elsewhere.[29] Provus outlined five stages of evaluation: definition, installation, process, product, and

---

[26]Egon Guba. "The Failure of Educational Evaluation." Educational Technology 9:29 (May 1969).

[27]Barbara Cass. "Application of Stufflebeam's Model to Large-Scale Program Evaluation." (Master's Thesis, University of Pittsburgh, 1968), p. 38.

[28]An excellent unpublished (July 1971) descriptive summary of different evaluation models which provides information in chart form has been prepared by James Sanders of the Educational Research and Evaluation Laboratory, Indiana University, Bloomington, Indiana.

[29]Esther Kresh. An Overview of the Discrepancy Model. (Pittsburgh: Office of Research, Pittsburgh Public Schools, 1969).

cost-benefit analysis. In its several years of operation in Pittsburgh, the evaluators never got beyond the third stage of the model in actual practice; and now that Provus has departed from Pittsburgh, the model is no longer used.

Conclusions

We have here briefly examined educational evaluation and more particularly curriculum evaluation. It has been seen that evaluation can be defined in many different ways and that there are both narrow and wide conceptions of the term. Several methodological approaches can be taken in educational evaluation including the PPBS, informal, and formal approaches. Educational evaluation models often stress different aspects of the curriculum process. For example, Stufflebeam has emphasized the decision-making element of evaluation. A number of distinctions concerning evaluation have been made which help to bring the evaluation process into sharper focus. These include the distinctions between formative and summative evaluation, between evaluation and analysis and between evaluation and description. A number of problems plague educational evaluation in addition to the problem of definition of terms. Evaluation designs, particularly experimental designs, often contain such flaws as lack of external validity and lack of a decision-making component. Inappropriate or insufficient criteria may result in faulty designs and reports. The generalizability of an evaluation is also a major problem with lack of random sampling, changes in courses and students across time, and the changing state of society presenting particular difficulties. Finally the problem of operationalization of evaluation models often stymies evaluators at the grass-roots levels.

Educational evaluation has made a great many advances in the last ten years, and its importance is becoming recognized at last. Despite these advances, it has been seen that a great many problems have yet to be overcome. These problems are now being examined in specific academic subject areas as well as on a more general scale. The problems enumerated in this brief overview of the evaluation of curricula must now be kept in mind as we move to a review of evaluation efforts in the Social Studies.

SOCIAL STUDIES EVALUATION: A REVIEW OF THE LITERATURE

The Problems

With the "traditional" History and Geography-centered Social Studies slowly being replaced in American schools, there has arisen a need for evaluation of the new Social Studies curricula. Problems impeding the development of evaluation include: (1) viewing of evaluation merely as testing and grading, (2) the limiting of evaluation to information collected by teachers, (3) inadequacies of standardized tests in Social Studies, (4) few and mostly descriptive evaluation models aimed specifically at Social Studies, and (5) models which disregard the important facet of decision-making. These and other problems, as seen in the current literature on Social Studies evaluation, will be examined here.

Evaluation done by teachers is often considered to be the only type of evaluation of importance. Unfortunately, this kind of evaluation usually only means "testing" and "grading." Evaluation by teachers, moreover, often refers simply to the measurement of student performance with no particular interest in the quality of the curriculum being taught. Some idea of the backwardness of this type of Social Studies evaluation can be obtained from a survey by Dwight Allen and Richard Gross in 1969 which reflected the use of the term "evaluation" as procedures specifically administered by teachers for testing purposes. Six hundred members of the National Council of the Social Studies, randomly selected from its mailing list, were asked to reply to 60 questions related to three elements of assessment: i.e., what should be covered in Social Studies evaluation; the administrative procedures for testing; and the kinds of examinations and test items needed. Some of the conclusions derived from the Allen and Gross survey are that:

1. Teachers frequently fail to relate their assessment practices to the aims they claim for their offerings.

2. Teachers are often inconsistent in their conception of evaluation.

3. Teachers are reticent to use the full range of evaluation techniques now available.

4. The use of many evaluation devices is misunderstood and such devices are often misused.

5. All of the purposes of evaluation are not understood by many teachers.

6. Teachers almost unanimously accept both essay and objective test items.

7. A disproportionate amount of time seems to be spent in the correction of English errors in social studies work.

8. The theory of sampling and test construction is not understood by teachers.

9. Few teachers employ item analyses or other checks upon their teaching, and their evaluation procedures.

10. Teachers, by their practices, encourage students to regard grading as a coercive weapon to be used against them.

11. Very few teachers perceive the major implications of the evaluation program which carry beyond the grading of students.[1]

This survey is interesting because it shows that Social Studies evaluation by teachers is indeed backward. But secondly, despite the fact that the Allen and Gross survey concerns "Problems and Practices in Social Studies Evaluation," the conclusions of the survey mostly seem to refer only to the role of the teacher in evaluation and this role was viewed essentially as testing and grading. The authors of the article from which the survey results were taken note that "evaluation should be used in a more comprehensive sense, in establishing realistic aims, in setting up purposeful curricula and courses, in critically reviewing instructional materials, and in judging the efficacy of our teaching process."[2] But their survey questions imply an unfortunately narrow view of evaluation. A broader survey of Social Studies evaluation is needed, one that will include evaluation procedures used in some of the major Social Studies curriculum projects, as well as "evaluation of evaluation," or meta-evaluation.

Gross and Allen do, nevertheless, note a few evaluation problems which they believe are special to Social Studies evaluation. Among them are:

---

[1] Dwight Allen and Richard Gross. "Problems and Practices in Social Studies Evaluation." in Teaching the Social Studies, by Richard Gross, W. McPhie, and Jack Fraenkel. (Scranton, Pa.: International Textbook Company, 1969), p. 480.

[2] Ibid.

1. The broad and imprecise goals that are held for the Social Studies...plus the future-oriented socio-civic purposes commonly expected to result from social education.

2. The problem of validity. Do the instruments we attempt to apply in evaluation really ascertain what we believe they do? Even as we build the necessary measures of skill we remain dogged by a doubt that has not been alleviated to date: are such tests of competency actually measures of the quality implied or are they largely indexes of ability to apply the knowledge already learned?

3. Among the major purposes of Social Studies goals are those reflecting desired socio-civic attitudes. Can a purely objective social science approach ensure progress in this domain?

4. The heart of the Social Studies program should rest in controversy. Often in controversy there is no answer or there is the possibility of a variety of hypotheses that either can't all be tested or that will not satisfy many who are involved.

5. Unhappily, the great bulk of commercial and standardized tests in the field of social studies have been found wanting by experts.[3]

Although Gross and Allen see the above problems as specific to Social Studies, the validity problem certainly is not nor is the problem of devising tests which measure quality instead of previously learned knowledge. And difficulties with standardized tests also plague other fields as well. This brings to light a major problem among writers on Social Studies evaluation: they do not seem to take note of what is happening in other fields as far as evaluation is concerned.

Something of the slow progress in Social Studies evaluation is evidenced in the 1965 Yearbook of the National Council of the Social Studies, "Evaluation in Social Studies." Where this anthology leans heavily toward "measurement" and "testing," a more recent yearbook, edited by Dorothy McClure Frazer, contains a more mature look at evaluation.

---

[3] Ibid., pp. 480-481.

As the Handbook of Research on Teaching has noted, research on evaluation, done especially in regard to Social Studies, over the last decade has been limited. "The need is acknowledged but the responsibility is left to professionals in evaluation."[4] Unfortunately, there has been little activity by professional evaluators in the area of Social Studies.

Social Studies Evaluation Models

Perhaps one of the most practical efforts at the development of an evaluation model for Social Studies education is that of William Stevens, Jr., and Irving Morrisset.[5] Called "A System for Analyzing Social Science Curricula," it offers a method for systematically analyzing curriculum materials and includes an extensive checklist of questions which users might ask of materials. Their outline has six headings: (1) Descriptive Characteristics, (2) Rationale and Objectives, (3) Antecedents Conditions, (4) Content, (5) Instructional Theory and Teaching Strategies, and (6) Overall Judgments. The Stevens and Morrisset model, being one of the few usable evaluation systems for the Social Studies, has been used more and more of late. Unfortunately, there is no concrete evidence concerning how well the model works. The authors have noted a few of the comments on their own model in an article in the EPIE Forum.[6] One of the criticisms that Stevens and Morrisset themselves take seriously is that their model is too unwieldy.

M. Frances Klein and Louise Tyler, in a critique of the Stevens and Morrisset model, however, find the complexity of the model as an asset. They see the system's value largely in the detailed comprehensiveness of the categories and questions formulated. They found that it "was the most comprehensive set of guidelines dealing with the analysis of curricula that the writers have seen." Klein and Tyler, moreover, note that several questions in the Stevens and Morrisset system suggest important standards for evaluating a social science curriculum:

1. Behavioral Objectives: Does the author word his specific objections in such a fashion that the verbs demonstrate student action-behavior that is clearly observable and/or measureable? Are specific guides to observation and measurement given?

---

[4] N. E. Gage. Handbook of Research on Teaching. (Chicago: Rand McNally, 1963), p. 1238.

[5] Dorothy McClure Frazer (ed.). Social Studies Curriculum Development. (Washington, D. C.: National Council of the Social Studies, 1969).

[6] William Stevens and Irving Morrissett. "A System for Analyzing Social Science Curricula." The EPIE Forum, 5 (December 1967), p. 15.

2. Articulation: Do the materials fit well with the existing curricula that wil precede and follow them? Do they fit well with materials in other subject studies simultaneously?[7]

The first questions on behavioral objectives, according to Tyler and Klein, infer a standard regarding the desireability of behaviorly defined objectives while the second set of questions on articulation suggests "a standard which would ensure the new materials were compatible and consistent with existing and future curricula."[8]

However, Tyler and Klein also point out a major flaw in the Stevens and Morrissett evaluation system. It is a flaw related to one of the distinctions discussed in the first part of this paper, namely that between description and evaluation, and it definitely limits the model. Stevens and Morrissett never state that their system is useful only for descriptive purposes. Tyler and Klein examine this flaw:

> Application of the set of taxonomic questions would result in a very extensive, detailed description of a curriculum... assuming that one could obtain answers to all or most of the questions. The results would not directly answer specific questions regarding what curricula should be. Many judgments might be inferred from the answers to the questions, but there is no specified criteria against which to formulate judgments regarding the quality of the curriculum.[9]

Scriven takes Stevens and Morrissett to task for some of the same reasons as Klein and Tyler... plus a few others. First of all, Scriven finds that the Curriculum Analysis System, in both its long and short forms, has "an absolutely pervasive sense of confusion about evaluation and description."

> Take, for example, this description under 2.2 (General Objectives): What are the generalized student outcomes that can be expected from the use of these materials? Obviously this calls for an evaluation and a crucial one.

---

[7] Frances Klein and Louise Tyler. "On Analyzing Curricula." Curriculum Theory Network, (Spring 1969), p. 23.

[8] Ibid.

[9] Ibid.

> It must be made clear whether this is what the producer says, or whether it is, in fact, part of the evaluation tucked into the Rationale section.[10]

Scriven also faults the Curriculum Analysis System for its commitment to the Bloom and Krathwohl taxonomy of objectives, which he notes has been shown to be seriously defective. He concludes that the Stevens and Morrissett system is "too committed to methodological and physical doctrines of excessive dubiety, excessively complicated, repetitious, and imprecisely described."[11] As this is the major evaluation system used in Social Studies education today, it would seem that Social Studies curriculum developers and evaluators are leaning on a "weak reed," and many do seem to use this system.

Aside from the Stevens-Morrissett system, the choice of evaluation systems specifically for Social Studies is very slim indeed. Dennis Gooler, in the Educational Product Report (October 1969), has taken Robert Stake's general evaluation model and tried to tie it into the evaluation of Social Studies curricula.[12] However, this model, too, as it is used by Gooler, is more descriptive than evaluative in nature. As Barbara Cass has indicated in her examination of seven evaluation systems, the Stufflebeam model (CIPP Model) should be much better in this regard.[13] Stake's model is also a good comprehensive descriptive model and should not be ignored. His distinctions among antecedent, transactional, and outcome data are very helpful. He defines antecedents as any conditions existing prior to teaching and learning which may relate to outcomes. Transactions are the encounters of students with teachers, student with student, author with reader, parent with counselor, ... the succession of engagements which comprise the process of education. Stake describes the outcome as a body of information including measurements of the impact of instruction on teachers, administrators, counselors and others, data on use of equipment, effects of the learning environment and cost incurred.[14]

---

[10] Michael Scriven. "An Introduction to Meta-Evaluation." Educational Product Report, 2 (February 1969), p. 38.

[11] Ibid.

[12] Dennis Gooler. "Evaluation and Change in the Social Studies." Educational Product Report, 3 (October 1969), pp. 6-13.

[13] Barbara Cass. Application of Stufflebeam's Model to Large-Scale Program Evaluation. (Master's Thesis, University of Pittsburgh, 1968), pp. 37-38.

[14] Robert Stake. "The Countenance of Educational Evaluation. Teachers College Record, 68 (April 1967), pp. 523-540.

Another and rather extensive look at social studies evaluation has been done by Roland F. Payette and C. Benjamin Cox in an article entitled, "New Dimensions in Evaluation of Social Studies Programs."[15] They also use the Stake model to assess evaluation in the social studies. Although the Cox and Payette work is detailed and quite comprehensive, it shares the Stake model's major flaw; i.e., a total neglect of curricular decision-making. Payette and Cox admit that their design will not take care of this critical area with the qualification that "it will not detail a step-by-step account of how to evaluate a social studies program as a means to making particular curricular decisions."[16]

Evaluation of Social Studies Materials

Other evaluation models in the Social Studies seem to aim particularly toward the evaluation of materials such as textbooks and equipment. These evaluations range from the quite systematic to the highly subjective. A prime example of bias in this type of "evaluation" is found in the Textbook Evaluation Reports issued by the Textbook Evaluation Committee of America's Future, Inc., from New Rochelle, New York. According to the "Committee," it came into being "because of much evidence of socialist and other propaganda in textbooks currently used in our secondary schools." They state that its purpose is:

> ...to evaluate these textbooks to determine the accuracy and competence and to report objectively the extent to which they give misleading or false impressions about our American form of government, our unique economy, our history and the relationship of the United States and its citizens to other countries and peoples of the world. The reviewers will also recommend textbooks which meet the Committee's standards.[17]

Unfortunately, it is a little difficult to tell what the Committee's standards are, since they never seem to be stated anywhere, but the political bent of the evaluations becomes obvious immediately. In the review of a social studies book (1955 vintage) called Living in Our Democracy, by Vanza Neulsen Devereaux and Homer Aker, the reviewer, Medford Evans (a former chief of training for the Atomic Energy Commission) begins by stating that:

---

[15] Roland Payette and Benjamin Cox. "New Dimensions in the Social Studies Programs." in Social Studies Curriculum Development: Prospects and Problems. ed. by Dorothy McClure Frazer. (Washington, D. C.: National Council of the Social Studies, 1969), p. 206.

[16] Ibid.

[17] Textbook Evaluation Reports. (New Rochelle, N.Y.: Textbook Evaluation Committee of America's Future, Inc., 1955), p. 1.

> ...this text is not bad as they go but it just shows how bad they go. It does not plug the Communist Party Line, as textbooks too often do, and indeed includes truly patriotic expressions; yet basically socialistic assumptions are taken for granted, so that following this test the conscientous teacher and pupil will plod ever deeper into socialist territory, where annexation by Communism may occur at almost any time.[18]

It is this type of biased review which has prompted Malorie Edelson to urge the production of "more objective and dependable textbook analyses."[19]

One attempt at the production of an objective evaluation instrument is the "Preliminary Instrument for Assessing Curriculum Materials, by Maurice J. Eash, which is quite detailed and comprehensive. The model is centered around four constructs:

1. Statement of Objectives
2. Organization of the material
3. Methodology (the modes of transaction used for engaging, focusing, and directing the learner)
4. Evaluation (guiding the learning through feedback as well as yielding data on the accomplishment of objectives.[20]

Descriptive in nature, this instrument should be most helpful to school boards and others who must choose textbooks for school systems. The Educational Product Report is also a valuable source in this regard.

A number of valuable guides to Social Studies curriculum projects are published by the Far West Laboratory for Educational Research and Development

---

[18] Ibid., p. 2.

[19] Malorie Edelson. "Evaluation and Confrontation: Keys to Textbook Change." Educational Product Report, 3 (November 1969), p. 6.

[20] Maurice Eash. "Assessing Curriculum Materials: A Preliminary Instrument." Educational Product Report, 2 (February 1969), p. 18.

in Berkeley, California. The reports concisely examine the various projects in terms of Goals and Objectives, Content and Materials, Classroom Strategy, Student and Teacher Prerequisites, Implementation, Requirements, Costs, Project History and Evaluation. Again these reports are descriptive in nature and quite helpful.

Conclusions

It can be seen that, outside the work of Taba and Stevens and Morrissett, little attention has been paid to the production of evaluation systems geared specifically to social studies. Robert Stake's model seems to have been adapted by several of the authors on social studies evaluation but the Stake model tends to ignore the important facet of decision-making. Most of the other materials related to evaluation in social studies are either purely "descriptive" or aimed at the assessment of social studies materials only. In other words, it must be concluded that evaluation, at least as seen from the literature specifically dealing with social studies evaluation, is a sad state of affairs indeed.

Is it necessary, however, to have an evaluation system aimed just at the social studies? Probably not. If social studies evaluators could use or modify other comprehensive models such as Stufflebeam's CIPP model, social studies evaluation may show marked improvement.

Having looked at the literature of Social Studies evaluation as it appears in recent books and journals, it must be noted that there is another very important source of evaluation literature, and this is from the Social Studies curriculum projects presently in operation in the United States. At this point, I would like to briefly examine evaluation as it is carried out in some of these projects.

EVALUATION AND THE SOCIAL STUDIES PROJECTS

There are between 30 and 40 Social Studies curriculum projects now in various stages of operation in the United States (see Appendix A). As Edwin Fenton has noted, the trickle of materials from these projects should soon reach flood proportions. Projects cover all aspects of the new Social Studies curriculum including history, geography, anthropology, economics, and sociology. Each of the projects, however, takes a different view of what evaluation should be. In this section of the paper, I will review evaluation problems and procedures carried out in nine different Social Studies projects currently underway.

Evaluation Problems in Curriculum Projects

Educators increasingly look to curriculum projects as the accepted medium for preparing classroom materials and for speeding up the innovative process in the curriculum. Projects are sponsored through universities and colleges, professional organizations such as the American Geographical Association, and non-profit educational organizations such as the Educational Development Corporation, Inc., in Cambridge, Massachusetts. Hulda Grobman has noted some of the problems which face curriculum development projects generally in an excellent AERA monograph.[1] First, as mentioned earlier, curriculum no longer is synonymous with the textbook and as curriculum projects have broadened their views of what curriculum includes, the scope of many of the evaluative investigations has also expanded.

Grobman also notes an often neglected aspect of evaluation in curriculum projects, that is the politics of evaluation, "...the extent to which systematic evaluation is politically feasible, that is the extent to which the project can afford the consequences of evaluation." In addition, there is the problem of the ethics of evaluation, a concern often ignored by writers on social studies evaluation. Information for evaluation often can be obtained either without the subject's knowledge that the information is being obtained or without his knowledge of the use to which it will be put. For example, some researchers have used "anonymous questionnaires" that are not really anonymous or interviews that are ostensibly to be used for one purpose when in actuality they will be used for another. The dubious justification for all this is that it is "effective."

---

[1] Hulda Grobman. Evaluation Activities of Curriculum Projects. (Chicago: Rand McNally, AERA Monograph No. 2, 1970), p. 3.

Grobman has observed that the kind of evaluation and the nature of the evaluation design are influenced by what is possible at a particular point in the project, when the project information will be used, the purpose for which it will be used, who will use it and the purpose of the whole project. With these points in mind, we may now turn to an examination of some of the social studies curriculum projects.

Evaluation in the Social Studies Projects: A Review

The Holt Social Studies Curriculum Project has been developed by Edwin Fenton of the Carnegie-Mellon University Social Studies Curriculum Center in Pittsburgh, Pennsylvania. The project began in 1963 and the materials referred to here were first used in 1964. However, the Social Studies Curriculum Center continues to produce a great number of new materials each year. The overall goal of this curriculum project, in the words of Fenton, was "to help each student develop to the limit of his ability into an independent thinker and a responsible citizen of a democracy." The program's specific objectives in terms of learning and development by students were: (a) the attainment of affective goals, the formation of attitudes which will encourage intellectual curiosity and independent inquiry, and the examination of values; (b) the acquisition of inquiry skills; and (c) the learning of content that is based on the latest scholarly knowledge.

The Holt (Carnegie-Mellon) Social Studies Curriculum has yet to be systematically evaluated. According to the Far West Laboratory for Educational Research and Development, which has reviewed the project,[2] the curriculum underwent classroom trial in five Pittsburgh schools and was evaluated on the basis of standardized tests and a Carnegie-designed Social Studies Inquiry Skills Test. In addition, a subjective questionnaire was used by the developers to record student evaluations. The Carnegie-Mellon staff briefly evaluated the first experimental courses for the ninth grade using STEP and Economic Understanding tests, but this evaluation, of a formative type, produced little usable data. It is somewhat interesting to note that under pressure from teachers who felt that the standardized tests were not measuring certain important skills, the Carnegie group designed their own Carnegie Test of Social Studies Inquiry Skills, which showed that experimental classes did better with the Carnegie materials. One might, however, be justifiably suspicious of an "in-house" developed test used to value in-house materials.

---

[2] Note: Descriptions of the nine projects in this chapter were taken from reports issued by the Far West Laboratory of Educational Research and Development, Berkeley, California.

Another study of the Holt-Carnegie materials, specifically the curriculum on Comparative Economic Systems, was evaluated by the NDEA Institute in History for Curriculum Specialists using the Stevens-Morrissett Curriculum Analysis scheme. Far West Laboratory notes that the Comparative Economics and the Comparative Political Science course are similar and that the conclusions of the Stevens-Morrissett model in this case may be the same:

1. The materials will probably be most useful to average and above average ability students. College-bound students would be most likely to achieve most of the objectives.

2. Below average or culturally deprived students, under the direction of a creative teacher, could achieve more objectives at a higher level of achievement through the audio-visual aspects of the course and through a conventional economics (or civics) course.

3. Minimum skills required of students are: ability to read at or near the ninth grade level, computational skills (for economics), and interpretive skills.

At this point the reader must be reminded of the heavy criticism of the Stevens-Morrissett model. As mentioned earlier, Scriven and Taylor and Klein have all showed that the scheme would produce descriptive results only, which it seems, judging from the vague conclusions above, is what happened when the model was applied to the Holt Curriculum. There does not seem to be any information on how this program compares with other programs or with any set of standards in terms of student effectiveness.

A second Social Studies project is that of the Educational Research Council (formerly called the Greater Cleveland Research Council) under the direction of Raymond English. The materials, once known as the Greater Cleveland Social Science Curriculum, will be published in final form by Allyn and Bacon. The Greater Cleveland Social Science Program began in 1961. The staff of this project voiced certain objections to traditional Social Studies programs such as the fact that traditional teaching relies too much on rote learning, that traditional curricula view Social Studies as bodies of content, that the teaching of Social Studies is too teacher-centered, and that texts are too fact-centered. Even after a decade's work it is still difficult to tell how well this curriculum, which attempts to remedy these complaints, will accomplish its objectives. By 1969 the new curriculum had been completed for the ninth grade level, and completion of the tenth, 11th, and 12th grade materials is scheduled for 1972. But, no formal evaluation procedures or tests of the program have been developed. Instead, the

developers have established a set of informal evaluative procedures, used in Cleveland area schools, which center around questions such as:

What concepts are too easy? Too difficult?

Is the program geared for the above average pupil? For all pupils?

What methods are most effective in developing the concepts?

Is the amount of material sufficient?

Is the material relevant?

So, despite a change of name, the project's evaluation system remains the same: vague, with subjective, informal attempts at formative evaluation and no firm plans for summative evaluation as yet. No comparisions with other curriculum projects have been made.

The Utah State University Social Studies Project is under the direction of James P. Shaver, who, like Fenton, is one of the leading lights of social studies curriculum development. Located at Utah State University at Logan, Utah, the project is funded by the U. S. Office of Education and materials resulting from the project will be published by Houghton-Mifflin Company. The initial proposal for the project was submitted by Dr. Shaver in 1965.

The developers set the following objectives for the project:

1. To develop an Outline of Concepts appropriate to and usable in the analysis of public issues.

2. To develop suggestions for teaching the concepts in the Outline.

3. To investigate interactions between different discussion styles and student personality traits as they affect learning of the concepts in the Outline.

The curriculum was designed for average to above-average students in grades 11 and 12 but the developers now state that the materials can be used at any secondary level.

The Utah materials were evaluated twice in two different Utah communities. The first field test involved 212 eleventh-grade students divided into four sections with two being taught by project staff and two by local non-staff

teachers. There was no control group. However, at the time of this first evaluation, the primary concern of Shaver's group was to study the relationship between teaching style and student personality, not evaluating the curriculum and its materials. Apparently, the project staff felt that, in their view at least, the evaluation of the earlier Harvard Social Studies Project (1964-67), which served as a model for the Utah Project, proved that their approach worked; (Shaver and Donald W. Oliver collaborated on the Harvard Project) and so, therefore, it was believed that further evaluation was not needed. The results of the second field test of the Utah materials have not as yet been reported. Even the reports of the assessment of the first field test were termed inconclusive. All in all, then, there is actually little evaluative evidence available about this project.

The Harvard Project Public Issues Series, which is now terminated, was directed by Donald W. Oliver, a Professor of Education at Harvard University, and was funded by the U. S. Office of Education. The Public Issues Series is published by American Education Publications and became available in the fall of 1970. The Harvard materials are designed for secondary-level students of average to above-average ability, and the teacher using the materials has to have substantial knowledge of controversial issues. In regard to specific objectives, the Harvard materials are designed "to teach students analytical skills that will enable them to discuss public issues more effectively." The content emphasis is on public issues and not on specific disciplines.

According to the Far West Laboratory for Educational Research and Development, the Harvard Project has carried out informal clinical evaluation of ongoing teaching and systematic evaluation of the program at its termination. No information seemed to be available on the informal evaluation, and the formal evaluation consisted of rating and content evaluation of oral discussion and "fairly conventional paper and pencil tests." Field tests of the Harvard Project materials were carried out in Newton, Massachusetts, High School from 1964 to 1967 with 46 students. Three control groups were used. The project evaluator has been quoted as stating that the evaluation design had two major weaknesses:

1. Students were not tested before or during the three-year program.

2. All testing was concentrated into a two-week period at the end of the school year.

The evaluator concluded that "the one-shot nature of the testing justifies some caution concerning the validity of the results." That would be an understatement, judging from the reports on evaluation in this project.

The Project staff constructed three written tests to measure learning outcomes relevant to the Project curriculum: a Concept Application Test, a structured Dialogue Analysis Test, and an Open-ended Dialogue Analysis Test. Also the Project developers constructed an open-ended test to compare students' retention of factual information about major American History topics. Additionally, the students took the standardized Educational Testing Service Problems of Democracy Test. The second part of the formal evaluation involved "dividing the students into groups of two and then asking them to discuss a fairly complicated case study they had read. Students were compared on the basis of these discussions. Although the results of these tests were quite favorable to the Project, the weaknesses mentioned above tend to destroy any sense of confidence one has in this partial evaluation.

The Law in American Society Project was developed by the Law in American Society Foundation under the direction of Robert H. Ratcliffe. The Project was jointly sponsored by the Chicago Board of Education and the Chicago Bar Association, and the materials became available from the Houghton-Mifflin Company in the fall of 1970. The curriculum is based on a series of materials entitled, Justice in Urban America, in which the emphasis is on how the law operates in the United States... what it really does in everyday life and what it can be made to do. The Project wants students to look at their problems in a lawyer-like way and to help them to deal with the problems of their communities as well as broad contemporary issues. The developers divided their specific objectives into four groups: (1) cognitive skills, (2) attitudes and values, (3) the process of inquiry, and (4) facts, principles, and generalizations from the discipline of law. The curriculum materials were designed for ninth graders but are believed to be suitable for all secondary students.

Little information seems to be available about the results of the evaluation which consisted of pre- and post-tests administered to ninth-grade classes who used the materials, in Chicago's Title I-supported schools and to a control group. The tests were designed to measure "students' growth in basic comprehension of the subject matter and attitude toward the function of the legal system, individual influence on the legal process, purposes of law enforcement, etc." In addition to these tests, other evaluation measures included classroom observation and questionnaires for teachers. The materials presently are being taught for the first time in a full-year course, but no information on results is available.

The Amherst Project's content emphasis is on American History, and the materials have been developed by the Committee on the Study of History, under the direction of Richard H. Brown. The curriculum is for any secondary grade level. The materials became available from the Addison-Wesley Publishing Company in the fall of 1970.

Field testing of each unit of the Project is determined by the number of teacher requests and by the number of copies available. Most of the units were tested in urban areas. The Project utilizes pre- and post-tests constructed by Drs. Rose Olver and David Schneider, who are psychologists at Amherst College. The tests attempt to measure "student control of major ideas and concepts" and include attitude scales, short-answer questions, semantic checklists, and reading-skill and interpretation questions. The Far West Laboratory notes that the tests are still in the developmental stage and that no hypotheses or criteria for field testing past or present have been published. In addition to this evaluation, teachers using the materials are required to complete daily logs and evaluate each unit to determine if it is clearly explained, if enough background material is provided and what might be added to make the teacher's manual more effective. Also, Amherst staff representatives visit every teacher using the materials at least once a year and report on classroom atmosphere, teacher's role and style, context of the class and lesson, and so forth. It is unfortunate that no published conclusions are available regarding this project's evaluation either.

The Lincoln-Filene Center Secondary Social Studies Program is designed for students of below-average to average ability levels. Its major content emphasis is on political science. Its developer and publisher is the Lincoln-Filene Center for Citizenship and Public Affairs, located at Tufts University, Medford, Massachusetts. The developers have two major objectives. The first is the affective objective of "helping each student increase his personal sense of political, social, and economic efficiency." The second objective is of a cognitive nature; that all citizens must learn the ways they can affect their environments. The curriculum materials, aimed at non-college bound students, also became available in the fall of 1970.

The evaluation of the narratives in the program is based on written tests and feedback from teachers and students. "The student materials are judged against three criteria: relevance to stated objectives; interest; and readability." The evaluation of the first criterion is based on objective tests (cognitive objectives), subjective tests (affective objectives), and informal feedback from students. "Interest" is determined by written evaluation and informal feedback from students, and "readability" is determined on the basis of content and vocabulary tests, evaluation by reading experts, and subjective observation by teachers.

Although there has been no "hard" evaluation of the program, subjective evaluation seems to show positive affective changes as a result of applying the curriculum.

Sociological Resources for the Social Studies is a project developed by the American Sociological Association and is under the direction of Robert C. Angell. The curriculum materials emphasize sociology and are aimed at

secondary students of average to above-average ability. The publisher of the materials is Allyn and Bacon, Inc. The project emphasizes involvement of the student in the process of inquiry, and its objectives, which are stated in general and not behavioral terms, concern such topics as the Role of Modern Women, Leadership in American Society and Analyzing Modern Organizations. It is anticipated that the materials will be completed about September 1971.

Evaluation of each unit or episode is done by means of various tests, interviews, questionnaires, and observation of trial classes. The tests used are a verbal ability test from the Psychological Corporation and a test designed by the project staff to measure students' mastery of the content of each episode. Teachers using the materials are given interviews and asked to complete questionnaires about an episode's effect. Classroom observation was also done. Field trials of the materials have occurred in Atlanta, Miami, Minneapolis, Seattle, and San Francisco. The Project does not yet have evaluation results for all of its episodes nor is there a comparison of the results of different episodes. Initial reaction to individual episodes is claimed to be favorable, however.

The last curriculum project to be reviewed is the High School Curriculum Center in Government Project developed by the High School Curriculum Center in Government at Indiana University in Bloomington, Indiana. The director of the project is Howard Mehlinger. The content emphasis is political science, and the materials, not as yet formally published, are designed for average to above-average secondary students. The developers' objectives concern the students' political knowledge, skills, and attitudes.

This project has developed one of the most comprehensive evaluation schemes of any of the above projects. John J. Patrick, co-author of the materials for this project with Howard Mehlinger, has written an excellent report on the evaluative procedures used. The project's materials were developed during a three-year period from 1966 to 1969. Two field trials of the materials have been conducted to date: during 1968-69 and 1969-70. Further evaluation study will be conducted during 1972-73, when the materials will be first used in published form. The first field study relied only upon feedback from questionnaires, classroom observation and students' test-answer sheets. The second evaluation study used more formal procedures, including a Political Knowledge Test, a Political Science Skills Test, and six political attitude scales, to measure student performance in terms of the instructional objectives of the course. The conclusions of this study appear quite favorable, and extensive rewriting of certain parts of the materials are taking place to correct defects discovered during the second field trial.

Conclusions

This cursory review of nine Social Studies projects now currently in operation across the United States covers materials published in 1970. Although the projects reviewed comprise only about one-fourth of the total number of ongoing projects, the sample does give some idea of the state of evaluation in all the projects. Results are generally disappointing and seem to bear out the lack of expertise in evaluation found in the more general literature of Social Studies evaluation. Only one project, the Holt program at Carnegie-Mellon University, seems to have used a fairly comprehensive and systematic form of evaluation. In this case, the Stevens-Morrissett System was used, and it is descriptive in nature and subject to the earlier-mentioned weaknesses. The Stevens-Morrissett scheme is reportedly being used in other projects also. Comprehensiveness in curriculum evaluation design definitely seems to be lacking. Many of the projects do not seem to be able to produce any meaningful results from the evaluations that have been done. At least one project's procedures were of highly dubious validity, and one has the feeling that several project staffs are groping in the dark when it comes to evaluative procedures. It should be noted in all fairness, however, that some of the data from the evaluations in projects has yet to be published. But, while some projects have not reached the stage where valid summative evaluation can be produced, the formative evaluation seems to be of poor quality and of little value. In addition, it seems that none of the projects have been compared with each other, and no standard criteria have emerged as yet which might aid comparative judgments. Assessments tend to be overly descriptive, the decision-making process is neglected, and informal and subjective evaluation procedures seem to be used far too regularly.

It could be concluded that Social Studies curriculum developers still have much to learn about evaluation. In many cases the only criteria of whether a project is "good" or not is how many books have been sold by the project's publisher. For many projects evaluation has definitely not risen above the level of "bookkeeping."

EVALUATION IN THE SOCIAL STUDIES: THE INTERNATIONAL PERSPECTIVE

The amount of information to be found about evaluation of Social Studies curricula in countries other than the United States is extremely slim. The information which is available, however, seems to indicate that the development of Social Studies and the evaluation of Social Studies curricula in other nations lags further behind developments in other subject areas than it does in this country. The improvements which have occurred have come about through the auspices of curriculum projects within a few nations and also through the efforts of international organizations. Here I shall briefly look at the work of these organizations and then at Social Studies and Social Studies evaluation in four countries: Britain, Australia, Canada, and India.

International Organizations and Social Studies Evaluation

An organization which works on the problems of curriculum development and evaluation on an international level is the International Association for the Evaluation of Educational Achievement (I. E. A.) based at the Wenner-Gren Center in Stockholm, Sweden. During the summer of 1971, the Association sponsored and administered a six-week International Seminar for Advanced Training in Curriculum Development and Innovation in Gränna, Sweden. Approximately 20 countries sent teams of six persons each to the Seminar to study such problems as the determination of the specifications for curriculum, the development of learning material and instructional procedures, the evaluation of the effectiveness of learning materials and instructional materials, and the in-service and pre-service training of teachers for curriculum changes. Benjamin Bloom directed the Seminar and also served as a member of the core faculty. Dr. Edwin Fenton served as the United States representative on Social Studies education.

UNESCO, another international organization, is also struggling with the problems of educational evaluation. UNESCO is trying to develop a "Standardized Data Reporting System" for the evaluation of its functional literacy programs. These programs, some of which have been in operation for three or more years, have never been evaluated. In addition, UNESCO has sponsored seminars on curriculum evaluation in several countries, the most widely reported of which was the seminar on the teaching of the social sciences at the secondary level at Burwood Teachers College in Victoria, Australia. The impact of this effort will be examined shortly.

Social Studies Evaluation in Other Nations

Among the individual nations which have made and continue to make some determined efforts at improving their Social Studies curricula are Great Britain, Australia, Canada, and India. In Great Britain, the Nuffield Humanities Project has shown concern for the promotion of curricula centered on controversial issues, but little evaluation information seems to be available. In the general area of curriculum evaluation, the British seem to be lagging behind the United States in their thinking on evaluation and curriculum development. Whitfield and Kerr found a marked absence in Britain "of examples in which trial pupil achievements have been measured in terms of specific curriculum objectives, an operation we term 'hard evaluation.'"[1]

Stephan Wiseman, in a review of British curriculum development and curriculum evaluation, expressed his anxiety over the apparent lack of enthusiasm of various educational organizations and committees for pursuing curriculum evaluation. He warned that:

> . . . evaluation is an essential part of the process of curriculum development: if it is ignored then the new syllabuses, produced with hope and optimism, will become entrenched as firmly as those they replace and the end result will be the substitution of a new set of claims for the imprisonment of the teachers and their pupils. The present wave of curriculum reform has occurred, somewhat belatedly, in response to demands of a rapidly changing environment. Such change will undoubtedly continue, and our new curricula must be capable of adpatation to continuing change.[2]

In Australia, too, concerted efforts are being made to improve Social Studies curricula. An important milestone in Social Studies improvement was the UNESCO seminar on the teaching of the social studies at the secondary level held at Burwood Teachers College in Victoria in 1967. D. G. Dufty, in a recent article, reviewed some of the curriculum changes which have occurred since the seminar. The Burwood seminar revealed widely diverging

---

[1] R. C. Whitfield and J. F. Kerr. "Some Problems in Course Evaluation." Teachers College Record, 72:267 (December 1970).

[2] Stephan Wiseman. "Curriculum Development and Curriculum Evaluation." Research in Education [Manchester University Press], 1:6 (May 1969).

views concerning the teaching of the social sciences, and it also revealed "the conservative nature of social science education in Australia compared with overseas countries."[3] The innovations which have occurred since Burwood have varied among the Australian states. The influence of American curriculum developers, particularly Fenton and Taba, appear in the changes made in some of the states. Dufty's article concludes, however, that social science curriculum development in Australia remains "an adoptive, adaptive, or eclectic process, and no attempt has yet been made to develop an entirely new model for a social science course, with the possible exception of the social science curriculum development at Monash University. Overseas visitors would find little here that is unique or well-researched."[4] Other writers indicate that educational evaluation procedures are generally undeveloped as well.

Canadian educators are showing increasing concern for evaluation as evidenced by the fact that in the fall of 1970, 251 educators enrolled in special two-day "evaluation research" clinics which travelled from Vancouver to Regina, Toronto, and Halifax under the auspices of the Canadian Educational Researchers' Association.[5] Each two-day session aimed to give a comprehensive picture of evaluation research through the presentation of papers by the clinic team as well as practice sessions which stressed methodology, statistical treatment of results, comparative methodologies, the analysis of current evaluation models, and the management of large projects via PERT. Garnet L. McDiarmid, writing on "The Meaning of Curriculum," has stressed that "evaluation of the curriculum becomes an evaluation of the general effect of the various inputs and not just an evaluation of a specific course."[6] McDiarmid also calls for clarification of terms, particularly the word "curriculum." Better clarification of the term "evaluation" is also needed, he feels. Norman France, writing in Education Canada on "Evaluation in the High School: Fact or Fiction,"[7] equates evaluation with testing, a confusion of terminology discussed earlier. There seems to be a little in the way of evaluation procedures aimed specifically at the Social Studies in Canada.

---

[3] D. G. Dufty. "After Burwood What? A Study in Curriculum Innovation and Evaluation." Australian Journal of Education, 15:75 (March 1971).

[4] Ibid., p. 93.

[5] Peggy Koopman. "What do We Mean by Evaluation Research." Education Canada, 11:48 (March 1971).

[6] Garnet McDiarmid. "The Meaning of Curriculum." Education Canada, 11:31 (March 1971).

[7] Norman France. "Evaluation in the High School: Fact or Fiction." Education Canada, 11:14-18 (March 1971).

The concept of Social Studies is still quite vague in India. P. K. Khasnavis has noted that Indian scholars talk about Social Studies, but it is doubtful if a common definition of Social Studies is shared by everyone. Indian educators are, nevertheless, now showing increasing interest in the field.[8] There is, for example, an informative little journal called Social Studies Teacher published by the University of Baroda. Moreover, a number of articles proposing reforms in Indian thinking about Social Studies have appeared in the journal in recent years. In addition, it would seem that Indian educators are beginning to look at American work in Social Studies. Since the importance of Social Studies as a subject is only beginning to be recognized, evaluation of Social Studies curricula is quite undeveloped. It is hoped that as the Indian educators develop Social Studies curricula, they will include evaluation as an integral part of the curricula from the outset.

It can be seen then that both Social Studies and Social Studies evaluation in other nations are considerable behind the work currently being done in the United States. The efforts of the international organizations such as UNESCO would seem to be one of the most important means of rapidly improving Social Studies in other countries. But it would also seem helpful if curriculum developers outside of the United States became better acquainted with the work of Provus, Guba, Stufflebeam, Stake, and other major American figures in the field of curriculum evaluation with the hope that evaluation in any subject area, including the Social Studies, could be rapidly improved.

---

[8] P. K. Khasnavis. "Meaning of Social Studies According to American Scholars." Social Studies Teacher [University of Baroda, India], 7:6 (July 1970).

CONCLUSION

Evaluation in the Social Studies is still, for the most part, at the level of bookkeeping. Most of the work on curriculum evaluation in the Social Studies has taken place in the United States, usually in curriculum projects financed from varying sources. Both the general literature on evaluation in the Social Studies and the evaluation efforts in the curriculum projects reveal the failures and weaknesses of curriculum evaluation in this subject area. Evaluation of Social Studies curricula in other nations remains considerably behind the work being done in the United States.

The Social Studies are changing and moving at long last after many years of stagnation. But if the field is to expand and develop in a systematic and intelligent manner, then a parallel and dynamic effort must also come about in Social Studies evaluation. That evaluation must not be something haphazardly "added on" to Social Studies development programs but must be an integral part of those programs from their inception. Perhaps, then, Social Studies will stop being the laggard among the disciplines and become more cohesive and worthwhile. "Social Studies," "Curriculum," and "Evaluation" are all difficult to define at the moment but evaluation may be the key to improvement of the curricula in the Social Studies as well as in other fields.

APPENDIX I

CURRICULUM DEVELOPMENT PROJECTS: SOCIAL STUDIES EDUCATION

I. Additional Social Studies Projects:

Asian Studies Curriculum Project
University of California at Berkeley

Committee on the Study of History
Amherst College, Amherst, Massachusetts

Economics Curricular Materials for Secondary Schools
Ohio University

Elementary Economics Project
University of Chicago, Industrial Relations Center

Elementary Social Science Education Program
University of Michigan, Ann Arbor, Michigan

Harvard University Social Studies Project
Harvard University, Cambridge, Massachusetts

Materials and Activities for Teachers and Children (MATCH)
Boston Children's Museum, Boston, Massachusetts

Our Working World
University of Colorado

Social Studies Curriculum Project
Carnegie-Mellon University, Pittsburgh, Pennsylvania

Sociological Resources for Secondary Schools
American Sociological Association

Taba Social Studies Curriculum Project
San Francisco State College, San Francisco, California

II. Project Directors:

Angell, Robert C.
Sociological Resources for the Social Studies

Brown, Richard H.
Amherst College

Collier, Malcolm
Anthropology Curriculum Study Project

Darcy, Robert L.
Ohio University

Dow, Peter
Education Development Center, Inc.

English, Raymond
Educational Research Council of America

Fenton, Edwin
Carnegie-Mellon University

Fox, Robert S.
University of Michigan

Halsey, Van R.
Amherst College

Kresse, Frederick H.
Boston Children's Museum

Lippitt, Ronald
University of Michigan

Lovenstein, Meno
Ohio University

Michaelis, John U.
University of California at Berkeley

Newmann, Fred M.
Harvard University

Oliver, Donald W.
Harvard University

Parsons, T. William
Anthropology Curriculum Study Project

Powell, Phillip E.
Ohio University

Rader, William D.
University of Chicago, Industrial Relations Center

Ratcliffe, Robert H.
Law in American Society Foundation

Senesh, Lawrence
University of Colorado

Wallen, Norman E.
San Francisco State College

III. Project Materials:

Four Communities Around the World (Grade 3)
San Francisco State College

From Subject to Citizen
Education Development Center, Inc.

Geography in an Urban Age
High School Geography Project

History as Culture Change: An Overview
Anthropology Curriculum Study Project

House of Ancient Greece
Boston Children's Museum

Humanities in Three Cities: An Inquiry Approach
Carnegie-Mellon University

Introduction to the Behavioral Sciences: An Inquiry Approach
Carnegie-Mellon University

The Japanese Family
Boston Children's Museum

Justice in Urban America Series
Law in American Society Foundation

Man: A Course of Study
Education Development Center, Inc.

Manpower and Economic Education: Opportunities in American Economic Life
Ohio University

Middle and South America: Societies in Transition (Grade 6)
San Francisco State College

Neighbors at Work
University of Colorado

New History of the United States: An Inquiry Approach
Carnegie-Mellon University

Our State: A Changing Society (Grade 4)
San Francisco State College

Readings in Sociology Series
Sociological Resources for the Social Studies

Shaping of Western Society: An Inquiry Approach
Carnegie-Mellon University

Social Science Laboratory Units
University of Michigan

Tradition and Change in Four Societies: An Inquiry Approach
Carnegie-Mellon University

United States and Canada: Societies in Transition (Grade 5)
San Francisco State College

United States: Change, Problems, and Promises (Grade 8)
San Francisco State College

Units in American History
Amherst College, The Committee on the Study of History

Western Civilization: Perspectives on Change (Grade 7)
San Francisco State College

World Studies Inquiry Series
University of California at Berkeley

APPENDIX II

EVALUATION: SOME DEFINITIONS

Alkin, Marvin C. "Evaluation of Theory Development." Evaluation Comment (October 1969), p. 1.

Evaluation - is the process of:

1. Ascertaining the decision areas of concern
2. Selecting appropriate information
3. Collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.

Evaluation Need Areas:

1. Systems Assessment
2. Program Planning
3. Program Implementation
4. Program Improvement
5. Program Certification

Astin, Alexander W. and Panos, Robert J. "The Evaluation of Educational Programs." in Educational Measurement (ed. by Robert L. Thorndike). Washington, D. C.: 1971, pp. 733.

Evaluation involves the collection of information concerning the impact of an educational program. An educational program is conceived as any ongoing educational activity which is designed to produce specified changes in the behavior of the individuals who are exposed to it. The major function of evaluation is to provide the decision maker with relevant information about the inputs, outputs, and operations of the program under construction.

Brooks, M. "The Community Action Program as a Setting for Applied Research." Journal of Applied Behavioral Science, No. 1 (1965), p. 337.

The procedure by which programs are studies to ascertain their effectiveness in the fulfillment of goals.

Caro, Francis G. "Issues in the Evaluation of Social Programs." Review of Educational Research, Vol. 41, No. 2 (April 1971).

Page 87:
...the procedure by which programs are studied to ascertain their effectiveness in the fulfillment of goals." (from B. G. Greenberg. "Evaluation of Social Programs." Review of the International Statistical Institute (1958), p. 36).

Page 88:

. . . methodological activity which combines performance data with a goal scale. (Michael Scriven. "The Methodology of Evaluation." AERA Monograph Series on Curriculum Evaluation No. 1, Perspectives of Curriculum Evaluation, 1967)

Cohen, David K. "Politics and Research: Evaluation of Social Action Programs in Education." Review of Educational Research, Vol. 40, No. 2 (1970), pp. 214, 232.

Evaluation is a mechanism with which the character of an educational enterprise can be explored and expressed. . . . Evaluation is a technique for measuring the satisfaction of public priorities; to evaluate a social action program is to establish an information system in which the main questions involve the allocation of power, status, and other public goods.

Cronbach, Lee J. "Course Improvement Through Evaluation." Teachers College Record, Vol. 64, No. 8 (May 1963), p. 672.

Evaluation may be defined as the collection and use of information to make decisions about an educational program.

Denny, Terry (ed.). "Educational Evaluation." Review of Educational Research, Vol. 40, No. 2 (April 1970), p. 181.

Evaluation--is the gathering of empirical evidence for decision-making and the justification of decision-making policies and the values upon which they are based. . . . The goal of evaluation must be to answer questions of selection, adoption, support and worth of educational materials and activities. . . . In the past we have avoided the goal of evaluation with its inherent threat to teachers, administrators, and curriculum developers and have concentrated on one or more of the non-threatening roles evaluation can play.

Denny, Terry. Educational Product Report, Vol. 2, No. 5 (February 1969).

Evaluation is a worrisome word in educational parlance that resists definition about as stoutly as any concept in vogue.

Doll, Ronald C. Curriculum Improvement. Boston: Allyn and Bacon, 1970, p. 379.

Evaluation may be defined as a broad and continuous effort to inquire into the effects of utilizing educational content and process according to clearly defined goals.

Glass, G. "The Growth of Evaluation Methodology." AERA Monograph Series on Curriculum Evaluation, No. 7, p. 2.

An attempt to assess the worth or social utility of a thing.

Gronlund, Norman E. Measurement and Evaluation of Teaching. New York: MacMillan Company, 1970, 2nd. ed.

Evaluation may be defined as a "systematic process of determining the extent to which educational objectives are achieved by pupils."

Guba, Egon G. Evaluation and Change in Education. A paper prepared for the Elk Grove Training and Development Center Spring Evaluation Conference, Arlington Heights, Illinois, May 16, 1968, p. 11.

Evaluation... is a process of providing and using information for making educational decisions... this definition sees evaluation as continuing (cyclical and looping), multi-faceted (involving many different methods and techniques), practical, and relevant.

Harris, Chester. "Some Issues in Evaluation." The Speech Teacher, 12: 191 (1963).

Evaluation is... the systematic attempt to gather evidence regarding student behavior that accompanies planned educational experiences.

Hayes, S. "Evaluating Development Projects." Paris: UNESCO, 1959, p. 16.

A body of concepts and practices which have proved their usefulness (in the field of social sciences) and which are applied in such a way that they can contribute to the improvement of practical activities.

Kresh, Esther. "An Overview of the Discrepancy Evaluation Model and a Related Case Study." *Presented at the Ohio Conference on Evaluation*, July 28-30, 1969. 26 p. (mimeo.)

Evaluation is the process of providing decision makers with relevant and timely information for making decisions.

Merriman, Howard O. Educational Evaluation and Decision-Making. New York: Phi Delta Kappa National Study Committee on Evaluation at the AERA Symposia on Educational Evaluation, February 1971, p. 1.

Definition: EDUCATIONAL Evaluation is the Process of Delineating, Obtaining and Providing Useful Information for Judging Decision Alternatives.

Key Terms:

Process - A particular and continuing activity subsuming many methods and involving a number of steps or operations.

Decision Alternatives - Two or more different actions that might be taken in response to some situation requiring altered action.

Information - Descriptive or interpretive data about entities (tangible or intangible) and their relationships, in terms of some purpose.

Delineating - Identifying evaluative information required through an inventory of the decision alternatives to be weighed and the criteria to be applied in weighing them.

Obtaining - Making information available through such processes as collecting, organizing, and analyzing and through such formal means as measurement, data processing, and statistical analysis.

Providing - Fitting information together into systems or sub-systems that best serve the purposes of the evaluation and reporting the information to the decision-maker.

Useful - Satisfying the scientific, practical, and prudential criteria (as specified in Chapter 1 of the Monograph) and pertaining to the judgmental criteria to be employed in choosing among the decision alternatives.

Judging - The act of choosing among several decision alternatives, the act of decision-making.

Scriven, Michael. "The Methodology of Evaluation." in Perspectives of Curriculum Evaluation. AERA Monograph Series No. 1, 1967, p. 43.

Formative Evaluation...is evaluation to improve a program during its course
Summative Evaluation...is evaluation designed to appraise a program after it is established.

Stills, David L. (ed.) "Evaluation Research." International Encyclopedia of the Social Sciences, Vol. 5. New York: The MacMillan Company and Free Press, 1968, p. 198.

The primary purpose of evaluation research is to "provide objective, systematic and comprehensive evidence on the degree to which the program achieves its intended objectives plus the degree to which it produces other unanticipated consequences"...

Stufflebeam, Daniel L. "Toward a Science of Educational Evaluation." Educational Technology, Vol. 8 (July 30, 1968), p. 5.

Evaluation means the provision of information through formal means, such as criteria, measurement and statistics, to serve as rational bases for making judgements in decision situations. A decision is a choice among alternatives. A decision situation is a set of alternatives. Judgment is the assignment of values to alternatives. A criterion is a rule by which values are assigned to alternatives, and optimally such a rule includes the specification of variables for measurement and standards for use in judging that which is measured. Statistics is the science of analyzing and interpreting sets of measurements. And, measurement is the assignment of numerals to entities according to rules, and such rules usually include the specification of sample elements, measuring devices and conditions for administering and scoring the measuring devices. Stated simply, evaluation is the science of providing information for decision-making.

Stufflebeam, Daniel L. "The Use and Abuse of Evaluation in Title III." p. 129.

More specifically, evaluation is defined herein as the process of acquiring and using information for making decisions associated with planning, programming, and recycling program activities.

Suchman, Edward. Evaluative Research. New York: Russell Sage Foundation, 1967.

Page 31:
...the determination (whether based on opinions, records, subjective, or objective data) of the results (whether desireable or undesireable; transient or permanent; immediate or delayed) attained by some activity(whether a program or part of a program, a drug or a therapy, an ongoing or a one-shot approach) designed to accomplish some valued goal or objective (whether ultimate, intermediate, or immediate, effort or performance, long or short range).

This definition contains four key dimensions: process--the "determination," criteria--the"results," stimulus--the activity, and value--the objective.

Page 31:
...we would like to propose a distinction between evaluation... as the general process of judging the worthwhileness of some activity regardless of the method employed, and evaluative research as the specific use of the scientific method for the purpose of making an evaluation.

Page 28:
...the process of determining the value or amount of success in achieving a predetermined objective. It includes at least the following steps: formation of the objective, identification of the proper criteria to be used in measuring success, determination and explanation of the degree of success, recommendations for further program activity. (from Glossary of Administrative Terms in Public Health, Vol. 50 (February 1960), pp. 225-226)

Page 28:
...the measurement of desirable and undesirable consequences of an action that has been taken in order to forward some goal we value. (from Riecken, Henry W. The Volunteer Work Camp: A Psychological Evaluation. Cambridge, Mass.: Addison-Wesley Press, 1952, p. 4)

Page 29:
...the procedures of fact-finding about the results of planned social action. (Hyman, Herbert H., et al. Applications of Methods of Education: Four Studies of the Encampment for Citizenship. Berkeley: University of California Press, 1962, p. 3)

Page 29:
Program evaluation can be defined as the measurement of success in reaching a stated objective. (from James, George. "Planning and Evaluation in Health Programs." in Administration of Community Health Services. Chicago: International City Managers Association, 1961, p. 124)

Page 30:
...a process which enables the administrator to describe the effects of his programme, and thereby to make progressive adjustments in order to reach his goals more effectively. (from Klineberg, Otto. The Problem of Evaluation." International Social Science Bulletin, Vol. 7, No. 3 (1955), pp. 346-352)

Page 11:
An evaluation is basically a judgment of worth--an appraisal of value

Taba, Hilda. Curriculum Development. New York: Harcourt, Brace, and World, 1962, p. 312.

Evaluation is the process of determining what these changes (in education) are, and appraising them against the values represented in objectives to find out how far the objectives in education are being achieved.

Tumin, Melvin M. "Evaluation of the Effectiveness of Education: Some Problems and Prospects." Interchange, Vol. 1, No. 3 (1970), p. 96.

Evaluation means different things to different people, not only because they are defensive about the possible results of a systematic scrutiny of their effectiveness, but also because they have different notions as to what are legitimate sources of pride and shame....Evaluation is a many-edged sword.

> It must unavoidably threaten and damage some of the actors in the scene, just as unavoidably it will improve the power position of others. The negative as well as the positive functions of evaluation must be seen and understood by the evaluator.... the process is one of discovery for the actors subject to evaluation and for the professionals engaged in the act. The latter must always be prepared to revise the initial models, for technical reasons, first, and because, second, they must, if they are at all sensitive, come to learn things about the connections among inputs, processes, and outcomes that they did not know before, no matter how experienced they may be.

Westbury, Ian. "Curriculum Evaluation." Review of Educational Research, Vol. 40, No. 2, p. 240.

> Curriculum evaluation is...a body of techniques, methodologies and principles created deliberately (and recently) to give some systematic form to the ways in which the assertion "we must evaluate" can be made to work.

BIBLIOGRAPHY

I. Books

Beatty, Walcott (ed.). Improving Educational Assessment and An Inventory of Measures of Affective Behavior. Washington, D. C.: Association for Supervision and Curriculum Development, 1969.

Curriculum Innovations and Evaluation. Princeton, N. J.: Educational Testing Service, 1969.

Doll, Ronald C. Curriculum Improvement. Boston: Allyn and Bacon, 1970.

Educational Evaluation: New Roles, New Means. Chicago: Sixty-Eighth Yearbook, National Society for the Study of Education, 1969.

Evaluative Research. Pittsburgh: American Institutes for Research, 1970.

Fancett, Verna. Social Science Concepts and the Classroom. Syracuse: Syracuse University Press, 1968.

Fenton, Edwin. The New School Studies. New York: Holt, Rinehart, and Winston, 1967.

Frazer, Dorothy McClure (ed.). Social Studies Curriculum Development: Problems and Prospects. Washington, D. C.: National Council for the Social Studies, 1969.

Gage, N. E. Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Grobman, Hulda. Evaluation Activities of Curriculum Projects. Chicago: Rand McNally (AERA Monograph), 1970.

Gross, Richard; McPhie, Walter; and Fraenkel, Jack. Teaching the Social Studies. Scranton, Pa.: International Textbook Company, 1969.

Guba, Egon, and Stufflebeam, Daniel L. Evaluation: The Process of Stimulating, Aiding, and Abetting Insightful Action. Columbus: Ohio State University, 1968.

_____. Strategies for the Institutionalization of the CIPP Evaluation Model. Columbus: Ohio State University, 1968.

International Encyclopedia of the Social Sciences. New York: Crowell, Collier, and MacMillan, Vol. 5, 1968.

On Evaluating Title I Programs. Princeton, N.J.: Educational Testing Service, 1966.

Price, Roy; Hickman, Warren; and Smith, Gerald. Major Concepts for the Social Studies. Syracuse: Syracuse University Press, 1965.

Social Studies Today: Guidelines for Curriculum Improvement. Harrisburg: Pennsylvania Department of Education, 1970.

Suchman, Edward. Evaluative Research. New York: Russell Sage Foundation, 1967.

Thomas, R. Murray; Sands, Lester; and Brubaker, Dale. Strategies of Curriculum Change. Scranton, Pa.: International Textbook Company, 1968.

Tyler, Ralph; Gagné, Robert; and Scriven, Michael. Perspectives of Curriculum Evaluation. Chicago: Rand McNally (AERA Monograph No. 1 on Curriculum Evaluation), 1967.

II. Articles

Alkin, Marvin C. "Educational Theory Development." Evaluation Comment, October 1969, pp. 2-7.

Anderson, C. Arnold. "A New Frame for the Social Sciences." Paper prepared for the NDEA Institute in History, 1966.

Baker, Robert L. "Curriculum Evaluation." Review of Educational Research, Vol. 39, June 1969, pp. 339-358.

Bassett, G. W. "Change in Australian Education." Australian Journal of Education, Vol. 15, March 1971, pp. 42-57.

Beller, E. Kuno. "Teacher Evaluation: Why, What, and How." Peabody Journal of Education, Vol. 48, January 1971, pp. 125-139.

Cahen, Leonard S. "Comparative Problems in Product and Curriculum Innovation." Educational Product Report, Vol. 2, February 1969, pp. 45-46.

Caro, Francis. "Approaches to Evaluative Research." Human Organization, Vol. 28, 28, Summer 1969, pp. 87-99.

Caro, Francis. "Issues in the Evaluation of Social Programs." Review of Educational Research, Vol. 41, April 1971, pp. 87-113.

Carpenter, Polly. "Analysis of Educational Programs." Santa Monica, Calif.: The Rand Corporation, 1971. (mimeo.)

Clawson, Elmer; Lowrie, Jane; and Macey, Joan. "Curriculum Materials Analysis I." Educational Product Report, Vol. 3, October 1969, pp. 20-31.

Cope, Robert G. "Observations in the Failure of Evaluation in Higher Education." The College of Education Record (University of Washington), Vol. 37, November 1970, pp. 3-6.

Cronbach, Lee J. "Course Improvement Through Evaluation." Teachers College Record, Vol. 64, May 1963, pp. 672-683.

Derr, C. Brooklyn. "Organizational Development and PPB for Education." Pittsburgh: University of Pittsburgh, 1971. (mimeo.)

Donnell, Nicole P. "EPIE Gathers User Data in Classroom Settings." Educational Product Report, Vol. 3, October 1969, pp. 14-19.

Dufty, D. G. "After Burwood, What? A Study in Curriculum Innovation and Evaluation." Australian Journal of Education, Vol. 15, March 1971, pp. 73-94.

Eash, Maurice J. "Assessing Curriculum Materials: A Preliminary Instrument." Educational Product Report, Vol. 2, February 1969, pp. 18-24.

Edelson, Malorie. "Evaluation and Confrontation: Keys to Textbook Change." Educational Product Report, Vol. 3, November 1969, pp. 6-11.

Enters, James, and Harms, Arden. "Curriculum Materials Analysis II." Educational Product Report, Vol. 3, October 1969, pp. 32-43.

"Evaluation Design." Tucson: Educational Innovators Press, 1970.

Fenton, Edwin. "The New Social Studies: Past, Present, and Future." Pittsburgh: Carnegie-Mellon University, 1970. (mimeo.)

Forehand, Garlie. "The Role of the Evaluator in Curriculum Research." Journal of Educational Measurement, Vol. 3, Fall 1966, pp. 199-204.

_____. "Functions of a Curriculum Evaluation System." Pittsburgh: Carnegie-Mellon University, n. d. (mimeo.)

France, Norman. "Evaluation in the High School: Fact or Fiction." Education Canada, Vol. 11, March 1971, pp. 14-18.

Freeman, Jack. "PPBS." Pittsburgh: University of Pittsburgh, 1971. (mimeo.)

Gooler, Dennis. "Evaluation and Change in the Social Studies." Educational Product Report, Vol. 3, October 1969, pp. 6-13.

Guba, Egon. "The Failure of Educational Evaluation." Educational Technology, Vol. 9, May 1969, pp. 29-38.

Hawkins, Colonel Taylor. "Educational Reform Through Planning Programming Budgeting Systems." Paper prepared for Special Training Workshop on Educational Finance and Development, International and Development Education Program, University of Pittsburgh, 1970. (mimeo.)

Hebbar, A. Sadanand. "An Experiment in Teaching Social Studies in Standard X." Social Studies Teacher (University of Baroda, India), Vol. 7, October 1970, pp. 18-26.

Johnson, Thomas J. "Evaluation in the Context of Product and Model Development in Laboratory and Research and Development Centers." St. Ann, Mo.: Central Midwestern Regional Educational Laboratory (CEMREL, Inc.), July 1971. (mimeo.)

Khasnavis, P. K. "Meaning of Social Studies According to American Educators." Social Studies Teacher (University of Baroda, India), Vol. 7, July 1970, pp. 6-9.

Klein, M. Frances, and Tyler, Louise. "On Analyzing Curricula." Curriculum Theory Network, Spring 1969, pp. 10-29.

Koopman, Peggy. "What Do We Mean by Evaluation Research?" Education Canada, Vol. 11, March 1971, pp. 48-51.

Kresh, Esther. "An Overview of the Discrepancy Evaluation Model and a Related Case Study." Pittsburgh, Pa.: Office of Research, Pittsburgh Public Schools, 1969.

Mason, Edwin. "Interdisciplinary Social Studies." The World and the School, Vol. 18, January 1970, pp. 90-94.

McDiarmid, Garnet L. "The Meaning of Curriculum" Education Canada, Vol. 11, March 1971, pp. 27-33.

McNaughton, A. F. "Curriculum as an Instrument of Change with Particular Reference to the Social Studies." Interchange, Vol. 1, 1970, pp. 61-69.

Meese, M. Kathryn. "A Model for Assessing Complex Educational Processes." Pittsburgh: University of Pittsburgh, 1971. (mimeo.)

Merriman, Howard O. "Educational Evaluation and Decision-Making." Paper presented at the AERA Conference, New York, February 1971.

Messick, Samuel. "Evaluation of Educational Programs as Research on the Educational Process." Paper presented at the American Psychological Association meetings, 1969.

Metcalf, Lawrence E., and Hunt, Maurice. "Relevance and the Curriculum." Phi Delta Kappan, Vol. 51, March 1970, pp. 358-361.

Patrick, John J. "The Impact of an Experimental Course, American Political Behavior, on the Knowledge, Skills and Attitudes of Secondary School Students." Bloomington, Ind.: Indiana University, 1971. (mimeo.)

_____. "American Political Behavior: A New Approach to Civic Education in Secondary Schools." Paper presented at NCSS Midwest Regional Conference on Teaching the New Social Studies, Des Moines, Iowa, March 22-23, 1971. (mimeo.)

Pulleybank, R. W. "Impediments to Change." Education Canada, Vol. 11, June 1971, pp. 13-17.

Sanders, James. "A Descriptive Summary of Frameworks for Planning Evaluation Studies." Bloomington, Ind.: Indiana University, July 1971. (mimeo.)

Scriven, Michael. "An Introduction to Meta-Evaluation." Educational Product Report, Vol. 2, February 1969, pp. 36-38.

Snyder, William. "Implementation of PPBS." Paper prepared for Special Training Workshop on Educational Finance and Development, International and Development Education Program, University of Pittsburgh, 1970. (mimeo.)

Stake, Robert E. "The Countenance of Educational Evaluation." Teachers College Record, Vol. 68, April 1967, pp. 523-540.

"Standardized Data Reporting System for Experimental World Literacy Programme." Paris: UNESCO, May 1971. (mimeo.)

Stevens, William, and Morrissett, Irving. "A System for Analyzing Social Science Curricula." The EPIE Forum, Vol. 1, No. 4, December 1967.

Stufflebeam, Daniel L. "Depth Study of the Evaluation Requirement." Theory into Practice, Vol. 5, June 1966, pp. 121-133.

_____. "Toward a Science of Educational Evaluation." Educational Technology, Vol. 8, July 30, 1968, pp. 5-12.

_____. "The Use and Abuse of Evaluation in Title III." Theory into Practice, Vol. 6, 1967, pp. 126-133.

Tom, Alan. "Curriculum Evaluation." in An Approach to Selecting Among Social Studies Curricula. St. Louis: St. Louis University, n. d. (mimeo.)

Tumin, Melvin. "Evaluation of the Effectiveness of Education: Some Problems and Prospects." Interchange, Vol. 1, 1970, pp. 96-109.

Tyler, Louise, and Klein, M. Francis. "Recommendations for Curriculum and Instruction Materials." Curriculum Theory Network, No. 1, Summer 1968.

Walberg, Herbert. "Curriculum Evaluation: Problems and Guidelines." Teachers College Record, Vol. 71, May 1970, pp. 557-570.

Wardrop, James. "Generalizability of Program Evaluation: The Danger of Limits." Educational Product Report, Vol. 2, February 1969, pp. 41-42.

Welch, Wayne. "Curriculum Evaluation." Review of Educational Research, Vol. 39, October 1969, pp. 429-443.

Westbury, Ian. "Curriculum Evaluation." Review of Educational Research, Vol. 40, April 1970, pp. 239-260.

Whitfield, R. D., and Kerr, J. F. "Some Problems in Course Evaluation: A Brithsh View." Teachers College Record, Vol. 72, December 1970, pp. 267-274.

Wiseman, Stephan. "Curriculum Development and Curriculum Evaluation." Research in Education (Manchester University Press, England), Vol. 1, May 1969, pp. 1-8.

Wood, Samuel. "A Multidimensional Model for the Observation, Analysis, and Assessment of Classroom Behavior." Morgantown, W. Va.: West Virginia University, 1970. (mimeo.)

Worthen, Blaine R.; Sanders, James; and Rogers, W. Todd. "Evaluation and Meta-Evaluation of the Eleventh Annual Phi Delta Kappa Research Symposium." Boulder: Laboratory of Educational Research, University of Colorado, 1971. (mimeo.)

ADDITIONAL MATERIALS

Taba, Hilda. Curriculum Development: Theory and Practice. New York: Harcourt, Brace, and World, 1962.

Reports: Far West Laboratory for Educational Research and development, Berkeley, California.

a. Holt Social Studies Curriculum Project
b. Educational Research Council Studies Program
c. Utah State University Social Studies Project
d. Harvard Project Public Issues Series
e. Law in American Society Project
f. Amherst Project
g. Lincoln-Filene Center Secondary Social Studies Program
h. Sociological Resources for the Social Studies
i. High School Curriculum Center in Government Project